MULTISCALE HIERARCHICAL DECOMPOSITION OF IMAGES WITH APPLICATIONS TO DEBLURRING, DENOISING AND SEGMENTATION

EITAN TADMOR, SUZANNE NEZZAR, AND LUMINITA VESE

ABSTRACT. We extend the ideas introduced in [TNV04] for hierarchical multiscale decompositions of images. Viewed as a function $f \in L^2(\Omega)$, a given image is hierarchically decomposed into the sum or product of simpler "atoms" u_k, where u_k extracts a more refined information from the previous scale u_{k-1}. To this end, the u_k's are obtained as dyadically scaled minimizers of standard functionals arising in image analysis. Thus, starting with $v_{-1} := f$ and letting v_k denote the residual at a given dyadic scale, $\lambda_k \sim 2^k$, then the recursive step $[u_k, v_k] = arginf \mathcal{Q}_T(v_{k-1}, \lambda_k)$ leads to the desired hierarchical decomposition, $f \sim \sum Tu_k$; here T is a blurring operator. We characterize such $\mathcal{Q}_T$-minimizers (by duality) and expand our previous energy estimates of the data f in terms of $\|u_k\|$. Numerical results illustrate applications of the new hierarchical multiscale decomposition for blurry images, images with additive and multiplicative noise and image segmentation.

CONTENTS

1. INTRODUCTION

We continue our study of the hierarchical image decomposition method introduced by the authors in [TNV04] (hereafter abbreviated TNV). We extend the hierarchical decomposition method to the case of functionals arising in image deblurring, in multiplicative image denoising, and in image segmentation. Convergence results and energy estimates are given, together with experimental results on real images.

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **04 NOV 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**
4. TITLE AND SUBTITLE **Multiscale Hierarchical Decomposition of Images with Applications to Deblurring, Denoising and Segmentation**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Unversity of Maryland,Center for Scientific Computation & Mathematical Modeling (CSCAMM),College park,MD,20472-3281**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We extend the ideas introduced in [TNV04] for hierarchical multiscale decompositions of images. Viewed as a function f 2 L2(), a given image is hierarchically decomposed into the sum or product of simpler "atoms" uk, where uk extracts a more refined information from the previous scale uk−1. To this end, the uk's are obtained as dyadically scaled minimizers of standard functionals arising in image analysis. Thus, starting with v−1 := f and letting vk denote the residual at a given dyadic scale, k 2k, then the recursive step [uk, vk] = arginfQT (vk−1, k) leads to the desired hierarchical decomposition, f PTuk; here T is a blurring operator. We characterize such QT -minimizers (by duality) and expand our previous energy estimates of the data f in terms of kukk. Numerical results illustrate applications of the new hierarchical multiscale decomposition for blurry images, images with additive and multiplicative noise and image segmentation.

15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT **Same as Report (SAR)**	18. NUMBER OF PAGES **27**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

1.1. **Hierarchical (X, Y) decompositions.** We begin with a pair of normed function spaces, $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$, and their associated $\mathcal{Q}$-functional,

$$(1.1) \qquad \mathcal{Q}(f, \lambda) := \mathcal{Q}(f, \lambda; X, Y) := \inf_{u \in X} \left\{ \|u\|_X + \lambda \|f - u\|_Y^2 \right\}, \quad \lambda > 0.$$

We will talk about functions in the smaller space X being "smoother" than those in Y, so that the condition $\|u\|_X < \infty$ can be viewed as a regularizing constraint. The use of regularized $\mathcal{Q}$-like functionals has a long history starting with the classical Tikhonov-type regularizations, consult [TA77, Mo84, Mo93]. They can be found in a variety of applications; we mention here two: the work on *support vector regression*, e.g., [Va98], and the early works in the context of *image processing* [GG84, AV94, RO94, GS00, GS00]. In these works, λ is treated as a fixed *threshold parameter*. The $\mathcal{Q}$-functional (1.1) is also closely related to the standard K-functional which arises in *interpolation theory*, e.g., [BL76, BS88, DL93, Kr07],

$$K(f, \lambda; X, Y) := \inf_{u \in X} \left\{ \|u\|_X + \lambda \|f - u\|_Y \right\}, \quad \lambda > 0.$$

The relation between the $\mathcal{Q}$- and the K-functionals is summarized in

$$\mathcal{Q}(f, \mu_\lambda; X, Y) \approx K(f, \lambda; X, Y), \quad \mu_\lambda := \frac{\lambda^2}{K(f, \lambda)}.$$

In this context of K-functionals, however, λ does not just serve as a threshold parameter, but in fact is treated as a *variable*: the collection of f's with prescribed behavior of $K(f, \lambda)$ as $\lambda \uparrow \infty$, forms intermediate interpolation smoothness spaces[*]. This is the point of view we adopt in the hierarchical decomposition of $\mathcal{Q}$-functionals described below, with λ being treated as a *scaling variable*.

We assume that the minimization problem has a solution $u := u_\lambda$, and we let v_λ denote the residual, $v_\lambda := f - u_\lambda$. This will be expressed as

$$(1.2a) \qquad f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] = \operatorname*{arginf}_{u+v=f} \mathcal{Q}(f, \lambda; X, Y).$$

In general, $\|u\|_X$ will be a regularizing term, thus u_λ will contain only the "larger" features of f, while the residual v_λ will contain the "smaller" features. Of course, the distinction between these two components is scale dependent – whatever is interpreted as 'small' features at a given λ-scale, may contain significant features when viewed under a refined scale, say 2λ,

$$(1.2b) \qquad v_\lambda = u_{2\lambda} + v_{2\lambda}, \quad [u_{2\lambda}, v_{2\lambda}] = \operatorname*{arginf}_{u+v=v_\lambda} \mathcal{Q}(v_\lambda, 2\lambda).$$

By combining (1.2a) with (1.2b) we arrive at a better two-scale representation of f given by $f \approx u_\lambda + u_{2\lambda}$. Features below scale $1/2\lambda$ remain unresolved in $v_{2\lambda}$ but, the process (1.2b) can be continued. Starting with an initial scale $\lambda = \lambda_0$,

$$(1.3a) \qquad f = u_0 + v_0, \quad [u_0, v_0] = \operatorname*{arginf}_{u+v=f} \mathcal{Q}(f, \lambda_0),$$

---

[*]Observe that the definition of the K-functional here exchanges the usual ordering between X and Y, so that it scales with *increasing* λ, in order to keep the compatability with the usual use of the scaling of $\mathcal{Q}$-functional in image processing.

a more refined decomposition of f into simpler "atoms" is obtained by successive application of the dyadic refinement step (1.2b),

$$(1.3b) \qquad v_j = u_{j+1} + v_{j+1}, \quad [u_{j+1}, v_{j+1}] := \operatorname*{arginf}_{u+v=v_j} \mathcal{Q}(v_j, \lambda_0 2^{j+1}; X, Y), \quad j = 0, 1, ...$$

After k such steps, we end up with the following *hierarchical decomposition* of f

$$
\begin{aligned}
f &= u_0 + v_0 = \\
&= u_0 + u_1 + v_1 = \\
&= \ldots\ldots \qquad = \\
(1.4) \qquad &= u_0 + u_1 + \cdots + u_k + v_k.
\end{aligned}
$$

The above multiscale expansion provides a new hierarchical representation of the data f,

$$(1.5) \qquad f \cong \sum_{j=0}^{\infty} u_j,$$

where the approximate equality $\cong$ in (1.5) should be interpreted as the convergence of the residuals v_k's in (1.4) to be made precise below. The partial sum, $\sum_{j=0}^{k} u_j$, provides a multi-layered description of f which lies in an intermediate scale of spaces, in between X and Y, though the precise regularity may vary, depending on the scales which are present in f.

Remark 1.1. [The homogeneity of the hierarchical decomposition]. We note here the anomaly of the $\mathcal{Q}$-functional (1.1): when an image f with minimizer $[u, v]$ doubles its intensity, $2f$, its minimizer does not scale accordingly since the quadratic-based $\mathcal{Q}$ is not homogeneous. This anomaly of the $\mathcal{Q}$-functionals is fixed by their hierarchical decompositions. To this end, we observe that if $[u_0, v_0]$ is the minimizing pair of $\mathcal{Q}(f, \lambda)$ then $[2u_0, 2v_0]$ is the minimizer of $\mathcal{Q}(2f, \lambda/2)$. Consequently, if an image f has the hierarchical description (1.3b), $f \cong \sum_{j=0}^{\infty} u_j$, then we find recursively,

$$2v_j = 2u_{j+1} + 2v_{j+1}, \quad [2u_{j+1}, 2v_{j+1}] := \operatorname*{arg\,inf}_{u+v=2v_j} \mathcal{Q}(v_j, \lambda_0 2^j; X, Y), \quad j = 0, 1, ...$$

We conclude that the hierarchical decomposition is homogeneous of degree one: when doubling the intensity, $2f$ has the corresponding hierarchical decomposition $2f \cong \sum_{j=0}^{\infty} 2u_j$.

1.2. Hierarchical (BV, L^2) decomposition. As a prototype example, we discuss the hierarchical (BV, L^2) decomposition introduced in TNV. A special case of Tikhonov regularization with $(X, Y) = (BV(\Omega), L^2(\Omega))$ is the Rudin-Osher-Fatemi Total Variation (TV) functional*. It was advocated by Rudin, Osher and Fatemi [ROF92] who proposed the functional, $\mathcal{Q}(f, \lambda, BV(\Omega), L^2(\Omega))$, to recover a sharp image u from its noisy version $f = u + v$, prescribed over an open bounded domain $\Omega \subset \mathbb{R}^2$. Here, $\|f - u\|_{L^2(\Omega)}^2$ is a fidelity term, $|u|_{BV(\Omega)} := \int |Du| = \|u\|_{BV(\Omega)}$ is a regularizing term preserving edges, and $\lambda > 0$ is a threshold parameter which measures their relative weight. For $f \in L^2(\Omega)$, the problem admits a unique minimizer $u := u_\lambda$ (consult [AV94], [CL97], or [Ve01] for a more general regularizing term), which decomposes an $L^2(\Omega)$-image, f, into two distinct components,

$$(1.6) \qquad f = u_\lambda + v_\lambda, \quad [u_\lambda, v_\lambda] = \operatorname*{arginf}_{u+v=f} \mathcal{Q}(f, \lambda; BV(\Omega), L^2(\Omega)).$$

---

*TV should not be confused with TNV.

This TV-based model is a very effective tool in denoising images while preserving edges provided apriori information on the noise scaling, λ, is known. Otherwise, if $\mathcal{Q}(f, \lambda)$ is being implemented with a too small λ, then only a cartoon representation of f is kept in the form of $u_\lambda \in BV(\Omega)$, while small textured patterns or oscillatory details are swept into the residual $v_\lambda := f - u_\lambda$. If, on the other hand, λ is kept too large, then u_λ remains loaded with too many details which is close to the original f; not much change has been applied to f and the compression ratio is small. In some cases, e.g., [ROF92, CL97], the parameter λ can be estimated if some statistical information on the noise is known, but in general, this setup is limited by the use of the one scale dictated by λ. The corresponding (BV, L^2)-hierarchical decomposition introduced by the authors in [TNV04],

$$(1.7) \qquad f \cong \sum_j u_j, \qquad u_j \in BV(\Omega),$$

is independent of apriori parameters and is particularly suitable for image representations. The resulting multi-layered expansion, $f \sim \sum_j u_j$, is *essentially nonlinear* in the sense that its dyadic blocks, u_j, depend on the data itself, $u_j = u_j(f)$. The dyadic blocks capture different layers of scales of the original image. Their precise multiscale nature is quantified in the energy decomposition derived in TNV,

$$\|f\|_{L^2(\Omega)}^2 = \sum_{j=0}^{\infty} \left[\frac{1}{\lambda_j} |u_j|_{BV(\Omega)} + \|u_j\|_{L^2(\Omega)}^2 \right].$$

Multi-layered representations of images are not new. We mention here those based on wavelet expansions, e.g., [MAC02], [ACMS98], and the TV based expansion suggested by Rudin and Caselles in [RC]. The hierarchical, multi-layered (BV, L^2)-decompositions (1.7) were found to be an effective tool in image processing, [TNV04, BCM05] and image registration [LPSX06, PL07]. They can be extended to application of image denoising in the presence of additive or multiplicative noise, to image deblurring or to image segmentation. These will be discussed in the sections below.

2. Hierarchical decomposition of blurry and noisy images

We are given a blurred image, represented by $f \in L^2(\Omega)$; blurring will be modeled by a linear, continuous blurring operator, $T : L^2(\Omega) \to L^2(\Omega)$ (such as a convolution with a Gaussian kernel). We consider a decomposition of f provided by the following $\mathcal{Q}_T(f, \lambda)$ minimization in the presence of blur,

$$(2.1a) \qquad \mathcal{Q}_T(f, \lambda; X(\Omega), L^2(\Omega)) := \inf_{u \in BV(\Omega)} \left\{ \|u\| + \lambda \|f - Tu\|_{L^2(\Omega)}^2 \right\}.$$

Here, the *regularization functional* $\| \cdot \| : X \subset BV(\Omega) \to [0, \infty]$ is a semi-norm which takes a general form,

$$(2.1b) \qquad \|u\| \equiv \|u\|_p := \int_\Omega \phi(D^p u), \qquad p \geq 1.$$

Few examples for such regularizing functionals are in order. If $\|u\| = |u|_{BV(\Omega)} = \int_\Omega |Du|$ is the total variation of u, then (2.1a) becomes the denoising-deblurring model introduced in [ROF92, RO94]. A more general BV-type model is provided by $\|u\| = \int_\Omega \phi(|Du|)$ with a proper norm, ϕ, defined on the space of measures; the corresponding $\mathcal{Q}_T$ in (2.1a) then becomes the generalized BV model studied in [DT84, Ve01, EO04]. Other examples, defined on dense

subspaces of BV, are provided by $\|u\| = \|Du\|_{L^2(\Omega)} : H^1(\Omega) \to [0, \infty)$, or the second-order $\|u\|_2 = \int_\Omega |D^2 u| : BH(\Omega) \to [0, \infty)$ defined on BH, the space of functions with bounded Hessian.

Let $v_\lambda := f - Tu_\lambda$ denote the 'texture' at scale λ associated with the blurring model (2.1a),

$$[u_\lambda, v_\lambda] = \arg\inf_{Tu+v=f} \mathcal{Q}_T(f, \lambda).$$

Starting with $\lambda = \lambda_0$ in (2.1a),

(2.2a) $$[u_0, v_0] = \arg\inf_{Tu+v=f} \mathcal{Q}_T(f, \lambda_0)$$

we proceed, by iterating at the dyadic scales $\lambda_j := \lambda_0 2^j$

(2.2b) $$[u_{j+1}, v_{j+1}] = \arg\inf_{Tu_{j+1}+v_{j+1}=v_j} \mathcal{Q}(v_j, \lambda_{j+1}), \qquad \lambda_j := \lambda_0 2^j, j = 0, 1, \ldots.$$

Thus, we have $v_j = Tu_{j+1} + v_{j+1}$ where $v_{-1} := f$. Summing the last recursive relation, we end up with hierarchical representation of the blurred image f

(2.3) $$f = Tu_0 + Tu_1 + \ldots + Tu_{k-1} + Tu_k + v_k,$$

which in turn paves the way for hierarchical, multiscale denoised expansion

(2.4) $$u \cong \sum_{j=0}^m u_j.$$

We note that the last expansion needs to be truncated at appropriately chosen finite m in order to avoid the ill-posedness which must occur as we accumulate infinitely many terms, approaching the "inversion" of the ill-conditioned T.

2.1. Hierarchical decomposition using $\mathcal{Q}_T$-minimizers.

To study the hierarchical expansions (2.4), we first characterize the minimizers of the $\mathcal{Q}_T$-functionals (2.1). The characterization summarized in theorem below extends Meyer's result [Me02, Theorem 4] (and we also refer to [ACM04, Chapter1] for related characterization of minimizers involving dual functionals).

We recall that the regularizing functional $\|f\|$ in (2.1b) is a semi-norm and we define its *dual* w.r.t the $L^2(\Omega)$ scalar product $\langle \cdot, \cdot \rangle$,

(2.5) $$\|f\|_* := \sup_{\|h\| \neq 0} \frac{\langle f, h \rangle}{\|h\|},$$

so that the usual duality holds

(2.6) $$\langle f, h \rangle \leq \|h\| \|f\|_*$$

We say that (f, h) is an *extremal pair* if equality holds above. The theorem below characterizes u as minimizer of the $\mathcal{Q}_T$-functional if and only if u and $T^*(f - Tu)$ form an extremal pair.

Theorem 2.1. *Let $T : L^2(\Omega) \to L^2(\Omega)$ be a linear continuous blurring operator with adjoint T^* and let $\mathcal{Q}_T$ denote the associated functional (2.1a).*

(i) The variational problem (2.1) admits a minimizer u. Moreover, if $\|\cdot\|$ is strictly convex, then a minimizer u with $\|u\| \neq 0$ is unique.

(ii) u is a minimizer of (2.1) if and only if

(2.7) $$\left\langle u, T^*(f - Tu) \right\rangle = \|u\| \cdot \|T^*f - T^*Tu\|_* = \frac{\|u\|}{2\lambda}.$$

The detailed proof of the theorem, whose second part was sketched in [Me02]), is postponed to the end of this section. The next two remarks lead to refinement of theorem 2.1, depending on the size of $\|T^*f\|_*$.

Remark 2.1. [The trivial minimizer]. More can be said in case f consists mostly of texture, in the sense that

(2.8) $$\|T^*f\|_* \leq \frac{1}{2\lambda}.$$

Indeed, if (2.8) holds, then the characterization of the minimizer u in (2.7) implies

$$\frac{\|u\|}{2\lambda} = \langle u, T^*f \rangle - \langle u, T^*Tu \rangle \leq \frac{\|u\|}{2\lambda} - \|Tu\|_{L^2}^2,$$

and hence $Tu = 0$. But then $\mathcal{Q}_T(f,\lambda) = \|u\| + \lambda\|f\|_{L^2}^2$ implies that $u \equiv 0$ is a minimizer of (2.1), with $\mathcal{Q}_T(f,\lambda) = \lambda\|f\|_{L^2}^2$. The converse of this assertion also holds. We summarize with the following corollary.

Corollary 2.2. *[The case $\|T^*f\|_* \leq 1/2\lambda$]. Let $T : L^2(\Omega) \to L^2(\Omega)$ be a linear continuous blurring operator with adjoint T^* and let $\mathcal{Q}_T$ denote the associated functional (2.1a). Then $\|T^*f\|_* \leq \frac{1}{2\lambda}$ if and only if $u \equiv 0$ is a minimizer of (2.1).*

Proof. Assume $\|T^*f\|_* \leq 1/2\lambda$. We have already seen that $u \equiv 0$ is a minimizer. One can also argue directly that since $2\lambda\langle T^*f, h \rangle \leq \|h\|$ for all $h \in BV(\Omega)$, then

$$\begin{aligned}
\|h\| + \lambda\|f - Th\|_{L^2(\Omega)}^2 &= \|h\| + \lambda\|f\|_{L^2(\Omega)}^2 - 2\lambda\langle f, Th \rangle + \lambda\|Th\|_{L^2(\Omega)}^2 \\
&= \|h\| - 2\lambda\langle T^*f, h \rangle + \lambda\|f\|_{L^2(\Omega)}^2 + \lambda\|Th\|_{L^2(\Omega)}^2 \\
&\geq \lambda\|f\|_{L^2(\Omega)}^2 + \lambda\|Th\|_{L^2(\Omega)}^2 \geq \lambda\|f - T0\|_{L^2(\Omega)}^2 + \|0\|,
\end{aligned}$$

and therefore $u \equiv 0$ is a minimizer of (2.1a). It remains to verify the "if" part, namely, if $u \equiv 0$ is a minimizer of (2.1), then for all $h \in BV(\Omega)$ we have,

(2.9) $$\lambda\|f\|_{L^2(\Omega)}^2 \leq \lambda\|Th - f\|_{L^2(\Omega)}^2 + \|h\|,$$

or

$$2\lambda\langle f, Th \rangle \leq \lambda\|Th\|_{L^2(\Omega)}^2 + \|h\|.$$

Rescaling $h \to \epsilon h$, we obtain

$$2\lambda\epsilon\langle f, Th \rangle \leq \lambda\epsilon^2\|Th\|_{L^2(\Omega)}^2 + \epsilon\|h\|.$$

Dividing by ϵ and letting $\epsilon \to 0_+$, yields $2\lambda\langle T^*f, h \rangle = 2\lambda\langle f, Th \rangle \leq \|h\|$ for all $h \in BV(\Omega)$ and we conclude $\|T^*f\|_* \leq 1/2\lambda$. $\square$

Remark 2.2. [Equivalence classes]. Consider the p-order semi-norm in (2.1b), $\|u\| \equiv \|u\|_p := \int_\Omega \phi(D^p u)$ and assume

$$\|T^*f\|_* < \infty.$$

We note that $\|\cdot\|_* \equiv \left(\|\cdot\|_p\right)_*$ should be considered on the complement of appropriate equivalence classes of "modulo polynomials of degree p". Indeed, since $\|h + P\|_p = \|h\|_p$ for any polynomial $P = P(x)$ of degree $\leq p - 1$, we have for arbitrary constant c,

$$\frac{\langle f, h + cP \rangle}{\|h + cP\|} = \frac{\langle f, h \rangle + \langle f, cP \rangle}{\|h\|} = \frac{\langle f, h \rangle}{\|h\|} + c\frac{\langle f, P \rangle}{\|h\|}.$$

Thus, $\|T^*f\|_* < \infty$, implies $\langle T^*f, P \rangle = 0$ for all $deg\, P \le p - 1$.

In particular, if we assume that

$$\frac{1}{2\lambda} < \|T^*f\|_* < \infty,$$

then a minimizer of $\mathcal{Q}_T(f)$ does not vanish, $\|u\| \ne 0$. Otherwise, if $\|u\| = 0$ then u is a polynomial of degree $\le p - 1$; but by the preceding argument, the polynomial u should be orthogonal to T^*f and hence

$$\mathcal{Q}_T(f, \lambda) = \lambda \|f - Tu\|_{L^2}^2 = \lambda \big(\|f\|_{L^2}^2 + \|Tu\|_{L^2}^2 \big),$$

which is minimized when $Tu = 0$. Given that $\|u\| = Tu = 0$, one can follow the proof of corollary 2.2, starting with (2.9) and concluding that $\|T^*f\|_* \le 1/2\lambda$, which contradicts our assumption. We can summarize this case in the following corollary.

Corollary 2.3. *[The case $\|T^*f\|_* > 1/2\lambda$]. Let $T : L^2(\Omega) \to L^2(\Omega)$ be a linear continuous blurring operator with adjoint T^* and let $\mathcal{Q}_T$ denote the associated functional (2.1a) with $\|\cdot\| = \|\cdot\|_p$. Assume that*

$$\frac{1}{2\lambda} < \|T^*f\|_* < \infty.$$

Then u is a minimizer of (2.1) if and only if u and $T^*(f - Tu)$ is an extremal pair and

$$(2.10) \qquad \|T^*f - T^*Tu\|_* = \frac{1}{2\lambda}.$$

Moreover, if $\|\cdot\|$ is strictly convex then the minimizer u is unique.

Proof. We can now divide the equality on the right of (2.7) by $\|u\| = \|u\|_p \ne 0$. Moreover, since $\|T^*u\|_* > 1/2\lambda$ then by remark 2.2, $\|u\| \ne 0$ and uniqueness follows from theorem 2.1(i). $\square$

Equipped with theorem 2.1, we can extend the (BV, L^2)-hierarchical decompositions introduced in TNV to general $\mathcal{Q}_T$-functionals.

Theorem 2.4.
(i) [Hierarchical expansion]. Let $f \in L^2(\Omega)$ and consider the dyadically-based $\mathcal{Q}_T$ decomposition (2.2). Then f admits the following hierarchical expansion

$$(2.11a) \qquad f \cong \sum_{j=0}^{\infty} Tu_j;$$

here, the $\cong$ should be interpreted as the convergence, $\sum_{j=0}^{k} Tu_j \to f$ in the weak $\|\cdot\|$-sense,

$$(2.11b) \qquad \left\|T^*\left(f - \sum_{j=0}^{k} Tu_j\right)\right\|_* = \frac{1}{\lambda_0 2^{k+1}}.$$

(ii) [Energy decomposition]. The following energy estimate holds

$$(2.12a) \qquad \sum_{j=0}^{\infty} \left[\frac{1}{\lambda_j}\|u_j\| + \|Tu_j\|_{L^2(\Omega)}^2\right] \le \|f\|_{L^2(\Omega)}^2, \quad \lambda_j := \lambda_0 2^j.$$

Moreover, if $f \in BV(\Omega)$ then equality holds in (2.12a),

(2.12b)
$$\sum_{j=0}^{\infty} \left[\frac{1}{\lambda_j} \|u_j\| + \|Tu_j\|_{L^2(\Omega)}^2 \right] = \|f\|_{L^2(\Omega)}^2, \quad f \in BV(\Omega).$$

We note in passing that the BV regularity assumption can be relaxed for the energy decomposition, *(2.12b), to hold; consult* [TNV04, Corollary 2.3].

Proof. If $\|T^*f\|_* \le 1/2\lambda$ then by corollary 2.2, the minimizer of (2.1a), $[u_\lambda, v_\lambda] = [0, f]$; otherwise

(2.13)
$$\|T^*v_\lambda\|_* = \frac{1}{2\lambda}, \quad \langle Tu_\lambda, v_\lambda \rangle = \frac{1}{2\lambda}\|u_\lambda\|.$$

The first statement (2.11a) then follows from the basic hierarchical expansion, $f = \sum_0^k Tu_j + v_k$ while noting that $\|T^*v_k\|_* = 1/2\lambda_k$. For the second statement, (2.12), we begin by squaring the basic refinement step, $Tu_{j+1} + v_{j+1} = v_j$,

(2.14) $$\|v_{j+1}\|_{L^2(\Omega)}^2 + \|Tu_{j+1}\|_{L^2(\Omega)}^2 + 2\langle Tu_{j+1}, v_{j+1} \rangle = \|v_j\|_{L^2(\Omega)}^2, \quad j = -1, 0, 1, \dots .$$

Observe that the last equality holds for $j = -1$ with v_{-1} interpreted as $v_{-1} := f$. We recall that $[u_{j+1}, v_{j+1}]$ is a minimizing pair for $\mathcal{Q}_T(v_j, \lambda_{j+1})$ and hence, by (2.13),

$$2\langle Tu_{j+1}, v_{j+1} \rangle = \frac{1}{\lambda_{j+1}}\|u_{j+1}\|,$$

yielding $\frac{1}{\lambda_{j+1}}\|u_{j+1}\| + \|Tu_{j+1}\|_{L^2(\Omega)}^2 = \|v_j\|_{L^2(\Omega)}^2 - \|v_{j+1}\|_{L^2(\Omega)}^2$. We sum up obtaining (2.12a)

$$\sum_{j=0}^{k} \left[\frac{1}{\lambda_j}\|u_j\| + \|Tu_j\|_{L^2(\Omega)}^2 \right] = \sum_{j=-1}^{k-1} \left[\frac{1}{\lambda_{j+1}}\|u_{j+1}\| + \|Tu_{j+1}\|_{L^2(\Omega)}^2 \right] =$$
$$= \|v_{-1}\|_{L^2(\Omega)}^2 - \|v_k\|_{L^2(\Omega)}^2 = \|f\|_{L^2(\Omega)}^2 - \|v_k\|_{L^2(\Omega)}^2 \le \|f\|_{L^2(\Omega)}^2.$$

Given that f has BV regularity, one can follow the argument in [TNV04, Theorem 2.2] to conclude the equality (2.12b). $\qquad\square$

We conclude this section with the promised

Proof. of theorem 2.1.

(i) The existence of a minimizer for the $\mathcal{Q}_T$-functional follows from standard arguments which we omit, consult e.g., [Jo07, Section 8.6]. We address the issue of uniqueness. Assume u_1 and u_2 are minimizers

$$\|u_i\| + \lambda\|f - Tu_i\|_{L^2}^2 = j_{min}, \quad i = 1, 2$$

then by standard line of argument we consider the average $u_3 := (u_1 + u_2)/2$ to find

$$\begin{aligned} j_{min} &\le \|u_3\| + \lambda\|f - Tu_3\|_{L^2}^2 \\ &\le \frac{1}{2}\left(\|u_1\| + \lambda\|f - Tu_1\|_{L^2}^2 + \|u_2\| + \lambda\|f - Tu_2\|_{L^2}^2 \right) - \frac{1}{4}\|T(u_1 - u_2)\|_{L^2}^2 \\ &\le j_{min} - \frac{1}{4}\|T(u_1 - u_2)\|_{L^2}^2. \end{aligned}$$

Thus, $u_1 - u_2$ belongs to the kernel of T. We then end up with the one-parameter family of minimizers, $u_\theta := u_1 + \theta(u_2 - u_1)$, $\theta \in [0, 1]$,

$$j_{min} \le \|u_\theta\| + \lambda\|f - Tu_\theta\|_{L^2}^2$$
$$\le \theta\|u_2\| + (1 - \theta)\|u_1\| + \theta\lambda\|f - Tu_2\|_{L^2}^2 + (1 - \theta)\lambda\|f - Tu_1\|_{L^2}^2 = j_{min}.$$

Clearly, the two minimizers satisfy $\|u_1\| = \|u_2\|$ and we conclude that the ball $\|u\| = \|u_1\| \ne 0$ contains the segment $\{u_\theta, \ \theta \in [0, 1]\}$, which by strict convexity, must be the trivial segment, i.e., $u_2 = u_1$.

(ii) If u is a minimizer of (2.1a), namely, if for any $h \in BV(\Omega)$ we have

$$\lambda\|f - T(u + \epsilon h)\|_{L^2(\Omega)}^2 + \|u + \epsilon h\| \ge \lambda\|f - Tu\|_{L^2(\Omega)}^2 + \|u\|,$$

then

(2.15) $$\qquad \lambda\epsilon^2\|Th\|_{L^2(\Omega)}^2 - 2\lambda\epsilon\langle Th, f - Tu\rangle + \|u + \epsilon h\| \ge \|u\|.$$

Since $\|\cdot\|$ is sublinear, the last inequality yields

$$\lambda\epsilon^2\|Th\|_{L^2(\Omega)}^2 - 2\lambda\epsilon\langle Th, f - Tu\rangle + \|u\| + |\epsilon|\|h\| \ge \|u\|,$$

or, after division by $\epsilon > 0$,

$$\|h\| + \lambda\epsilon\|Th\|_{L^2(\Omega)}^2 \ge 2\lambda\langle Th, f - Tu\rangle.$$

Letting $\epsilon \downarrow 0_+$, we obtain $\|h\| \ge 2\lambda\langle Th, f - Tu\rangle$ for any $h \in BV(\Omega)$ and hence

(2.16) $$\qquad \frac{1}{2\lambda} \ge \sup_{\|h\| \ne 0} \frac{\langle h, T^*(f - Tu)\rangle}{\|h\|} = \|T^*f - T^*Tu\|_*.$$

To confirm (2.7) it remains verify the reverse inequality in (2.16). To this end we set $h = u$ and $-1 < \epsilon < 0$ in (2.15), yielding

$$\lambda\epsilon^2\|Tu\|_{L^2(\Omega)}^2 + (1 + \epsilon)\|u\| \ge \|u\| + 2\lambda\epsilon\langle u, T^*(f - Tu)\rangle,$$

or

$$\lambda\epsilon^2\|Tu\|_{L^2(\Omega)}^2 + \epsilon\|u\| \ge 2\lambda\epsilon\langle u, T^*(f - Tu)\rangle.$$

Dividing by ϵ and letting $\epsilon \uparrow 0_-$, we obtain $\frac{1}{2\lambda}\|u\| \le \langle u, T^*(f - Tu)\rangle$. This, together with (2.16) implies,

$$\frac{1}{2\lambda}\|u\| \le \langle u, T^*(f - Tu)\rangle \le \|u\|\|T^*f - T^*Tu\|_* \le \frac{1}{2\lambda}\|u\|.$$

so that the last inequalities become equalities and (2.7) follows.

Conversely, we show that if u satisfies (2.7) then it is the desired minimizer. To this end, we rewrite

$$\|f - T(u + h)\|_{L^2(\Omega)}^2 \equiv \big\langle (f - Tu) - Th, f - T(u + h)\big\rangle$$
$$\equiv \big\langle f - Tu, f + T(u + h)\big\rangle - 2\big\langle f - Tu, T(u + h)\big\rangle - \big\langle Th, f - Tu\big\rangle + \|Th\|_{L^2(\Omega)}^2$$
$$\equiv \|f - Tu\|_{L^2(\Omega)}^2 - 2\overbrace{\big\langle f - Tu, T(u + h)\big\rangle}^{\#1} + 2\overbrace{\big\langle Tu, f - Tu\big\rangle}^{\#2} + \|Th\|_{L^2(\Omega)}^2.$$

Now, assumption (2.7), $\|T^*(f - Tu)\|_* = \frac{1}{2\lambda}$, implies $2\lambda(\#1) \leq \|u + h\|$ and $2\lambda(\#2) = \|u\|$. We conclude that for any $h \in BV(\Omega)$,

$$\|u + h\| + \lambda\|f - T(u + h)\|^2_{L^2(\Omega)}$$
$$= \|u + h\| - 2\lambda(\#1) + 2\lambda(\#2) + \lambda\|f - Tu\|^2_{L^2(\Omega)} + \lambda\|Th\|^2_{L^2(\Omega)} \geq \|u\| + \lambda\|f - Tu\|^2_{L^2(\Omega)}.$$

Thus, u is a minimizer of (2.1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Remark that a lack of uniqueness is demonstrated in an example of [Me02, pp. 40], using the ℓ^∞-unit ball, which in turn lacks strict convexity. Thus, strict convexity is necessary and sufficient for uniqueness.

2.2. Discretization of Euler-Lagrange and numerical results.

We consider for illustration the case of the total variation [RO94], therefore $\|u\| = \int_\Omega |Du|$. In practice, we simplify the formulation by working only on $W^{1,1}(\Omega)$ and we write $\|u\| = \int_\Omega |\nabla u| dx dy$. In order to construct the hierarchical decomposition of f, we use the associated Euler-Lagrange equation of $\mathcal{Q}(f, \lambda)$

$$T^*Tu_\lambda = T^*f + \frac{1}{2\lambda} \text{ div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|}\right).$$

When working on a bounded domain Ω, we augment the Euler-Lagrange equations by the following Neumann boundary condition:

$$\frac{\partial u_\lambda}{\partial n}|_{\partial\Omega} = 0.$$

The hierarchical decomposition, $f \sim \sum_{j=0}^k Tu_j = T\sum_{j=0}^k u_j$, is obtained. Note that we are really interested in the deblurred image $u = \sum_{j=0}^k u_j$, from which the u_j's are constructed as (approximate) solutions of the recursive relation governed by the following PDE:

$$T^*Tu_{j+1} - \frac{1}{2\lambda_{j+1}} \text{ div} \left(\frac{\nabla u_{j+1}}{|\nabla u_{j+1}|}\right) = \overbrace{-\frac{1}{2\lambda_j} \text{ div} \left(\frac{\nabla u_j}{|\nabla u_j|}\right)}^{T^*v_j}.$$

We implement our algorithm for this type of image in essentially the same way as for the case without blurring (see [TNV04]). The only difference is that we have to deal with the blurring operator T, a Gaussian kernel in our experiments, and in this we follow the method of discretization in [AV97] and [Ve01]. The first step is to remove the singularity when $|\nabla u_\lambda| = 0$, by replacing $\mathcal{Q}_T(f, \lambda)$ with

$$\mathcal{Q}_T^\epsilon(f, \lambda) := \inf_{u \in BV} \left\{ \int_\Omega \sqrt{\epsilon^2 + |\nabla u|^2} dx dy + \lambda\|f - Tu\|^2_{L^2(\Omega)}\right\}.$$

We find the minimizer, $u_\lambda \equiv u_{\lambda,\epsilon}$ of the regularized functional associated with $\mathcal{Q}_T^\epsilon$, at each step of our hierarchical decomposition. The associated Euler-Lagrange equations are

$$(2.17a) \qquad T^*Tu_\lambda = T^*f + \frac{1}{2\lambda} \text{ div} \left(\frac{\nabla u_\lambda}{\sqrt{\epsilon^2 + |\nabla u_\lambda|^2}}\right) \text{ in } \Omega,$$

$$(2.17b) \qquad\qquad\qquad\qquad \frac{\partial u_\lambda}{\partial n} = 0 \text{ on } \partial\Omega,$$

that we solve by a dynamic gradient descent scheme $(x, y, t) \mapsto u(x, y, t)$,

(2.18)
$$\frac{\partial u}{\partial t} + T^*Tu_\lambda = T^*f + \frac{1}{2\lambda} \text{div} \left(\frac{\nabla u_\lambda}{\sqrt{\epsilon^2 + |\nabla u_\lambda|^2}} \right) \text{ in } \Omega \times [0, +\infty), \quad u(x, y, 0) = f(x, y) \text{ in } \Omega.$$

As in TNV, we use a computational grid, $(x_i := ih, y_j := jh, t^n := n\triangle t)$, to cover the domain Ω for $t \geq 0$, where h is the cell size. Let $D_+ = D_+(h)$, $D_- = D_-(h)$ and $D_0 := (D_+ + D_-)/2$ denote the usual forward, backward and centered divided difference. As an example of differencing in the x-direction, $D_+^x u_{i,j} = (u_{i+1,j} - u_{i,j})/h$.

We discretize (2.18) as follows

$$\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\triangle t} + T^*Tu_{i,j}^n = T^*f_{i,j} + \frac{1}{2\lambda}D_-^x \left[\frac{(u_{i+1,j}^n - u_{i,j}^{n+1})/h}{\sqrt{\epsilon^2 + (D_+^x u_{i,j}^n)^2 + (D_+^y u_{i,j}^n)^2}} \right]$$

$$+ \frac{1}{2\lambda}D_-^y \left[\frac{(u_{i,j+1}^n - u_{i,j}^{n+1})/h}{\sqrt{\epsilon^2 + (D_+^x u_{i,j}^n)^2 + (D_+^y u_{i,j}^n)^2}} \right].$$

Following [AV97, Ve01], we will work with convolution-type T's which are realized by symmetric matrices, $(C_{\alpha\beta})_{\alpha,\beta=1,d}$,

$$Tu_{i,j} := \sum_{\alpha,\beta=1}^{d} C_{\alpha\beta}u_{i+d/2-\alpha,j+d/2-\beta}, \qquad \sum_{\alpha,\beta=1}^{d} C_{\alpha\beta} = 1.$$

Since T is assumed to be symmetric, T^*Tu is approximated by

$$T^2 u_{i,j} = \sum_{\alpha,\beta=1}^{d} \sum_{\gamma,\delta=1}^{d} C_{\alpha\beta}C_{\gamma\delta}u_{i+d-\alpha-\gamma,j+d-\beta-\delta}.$$

To implement our decomposition, either a fixed point Jacobi or Gauss-Seidel iterative method can be used to solve the discrete regularized Euler-Lagrange equations above. Introducing the notations:

(2.19a)
$$c_E(u_{i,j}^n) = \frac{1}{\sqrt{\epsilon^2 + \left(\frac{u_{i+1,j}^n - u_{i,j}^n}{h} \right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j}^n}{h} \right)^2}},$$

(2.19b)
$$c_W(u_{i,j}^n) = \frac{1}{\sqrt{\epsilon^2 + \left(\frac{u_{i,j}^n - u_{i-1,j}^n}{h} \right)^2 + \left(\frac{u_{i-1,j+1}^n - u_{i-1,j}^n}{h} \right)^2}},$$

(2.19c)
$$c_S(u_{i,j}^n) = \frac{1}{\sqrt{\epsilon^2 + \left(\frac{u_{i+1,j}^n - u_{i,j}^n}{h} \right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j}^n}{h} \right)^2}},$$

(2.19d)
$$c_N(u_{i,j}^n) = \frac{1}{\sqrt{\epsilon^2 + \left(\frac{u_{i+1,j-1}^n - u_{i,j-1}^n}{h} \right)^2 + \left(\frac{u_{i,j}^n - u_{i,j-1}^n}{h} \right)^2}},$$

we have

$$u_{i,j}^{n+1} + 2\triangle t \lambda h^2 T^* T u_{i,j}^n + \triangle t \Big(c_E(u_{i,j}^n) + c_W(u_{i,j}^n) + c_S(u_{i,j}^n) + c_N(u_{i,j}^n) \Big) u_{i,j}^{n+1}$$

$$= u_{i,j}^n + \triangle t \Big(c_E(u_{i,j}^n) u_{i+1,j}^n + c_W(u_{i,j}^n) u_{i-1,j}^n + c_S(u_{i,j}^n) u_{i,j+1}^n + c_N(u_{i,j}^n) u_{i,j-1}^n + 2\lambda h^2 T^* f_{i,j} \Big).$$

By using the most recent values of the $u_{i,j}$'s, we implement the Gauss-Seidel scheme.

In Figure 2.1, we show a test image, its blurry version (with Gaussian blur, no noise), the results obtained using the Rudin-Osher model [RO94], and the results obtained using the hierarchical model. The proposed hierarchical model gives improved results over the standard RO model, in terms of the root mean square error,

$$rmse := \frac{\sqrt{\sum_{i,j}(u_{orig,i,j} - u_{restored,i,j})^2}}{\sum_{i,j} 1}.$$

Similarly Figure 2.2 illustrates how the hierarchical decomposition works for noisy blurred images, again as an improvement over the Rudin-Osher model.

3. THE HIERARCHICAL (SBV, L^2) DECOMPOSITION

We want to construct the hierarchical decomposition based on the Mumford-Shah functional [MS89]. To this end we consider its elliptic approximation of Ambrosio and Tortorelli [AT92],

$$\mathcal{AT}_\rho(f, \lambda) := \inf_{\{u,v,w \ | \ u+v=f\}} \left\{ \mu \sqrt{\int_\Omega (w^2 + \rho \epsilon_\rho) |\nabla u|^2 dx} + \rho \|\nabla w\|_{L^2(\Omega)}^2 + \frac{\|w - 1\|_{L^2(\Omega)}^2}{4\rho} + \lambda \|v\|_{L^2(\Omega)}^2 \right\},$$

where $\epsilon_\rho \to 0$ as $\rho \downarrow 0$ and λ, μ are positive weight parameters.

Remark 3.1. Note that we modified the $\mathcal{AT}_\rho$-functional, where the square-root first-term on the right is replacing the original term, $\int_\Omega (w^2 + \rho \epsilon_\rho) |\nabla u|^2 dx$, appearing in Ambrosio-Tortorelli work [AT92]. Our modified $\mathcal{AT}_\rho$-functional does not affect the main properties of the segmentation model, however; it is introduced here in order to enable the characterization of $\mathcal{AT}_\rho$-minimizers in section 3.1 below. Our numerical calculations will then utilize the original formulation $\mathcal{AT}_\rho$-functional.

Let $[u_\lambda, v_\lambda]$ be the minimizer of $\mathcal{AT}_\rho(f, \lambda)$ (depending on w). Here $f \in L^\infty(\Omega)$ and u_λ is restricted to the smaller SBV space (— a special subclass of BV space, consisting of measure gradients free of the Cantor component, [AT92]), while the texture v_λ lives in L^2. We proceed to construct the hierarchical (SBV, L^2) decomposition of f in the same manner as before, letting $[u_{j+1}, v_{j+1}]$ be the AT minimizer

$$[u_{j+1}, v_{j+1}] = \arg \inf_{u,v,w,u+v=v_j} \mathcal{AT}_\rho(v_j, \lambda_j), \quad \lambda_j = \lambda_0 2^j.$$

We end up with the hierarchical decomposition

$$f = u_0 + u_1 + ... + u_k + v_k.$$

Here, at each hierarchical step, we also obtain the *edge detectors* $1 - w_j = 1 - w_{\lambda_j}$, which are (essentially) supported along the boundaries of objects enclosed by edges identified by u_j.
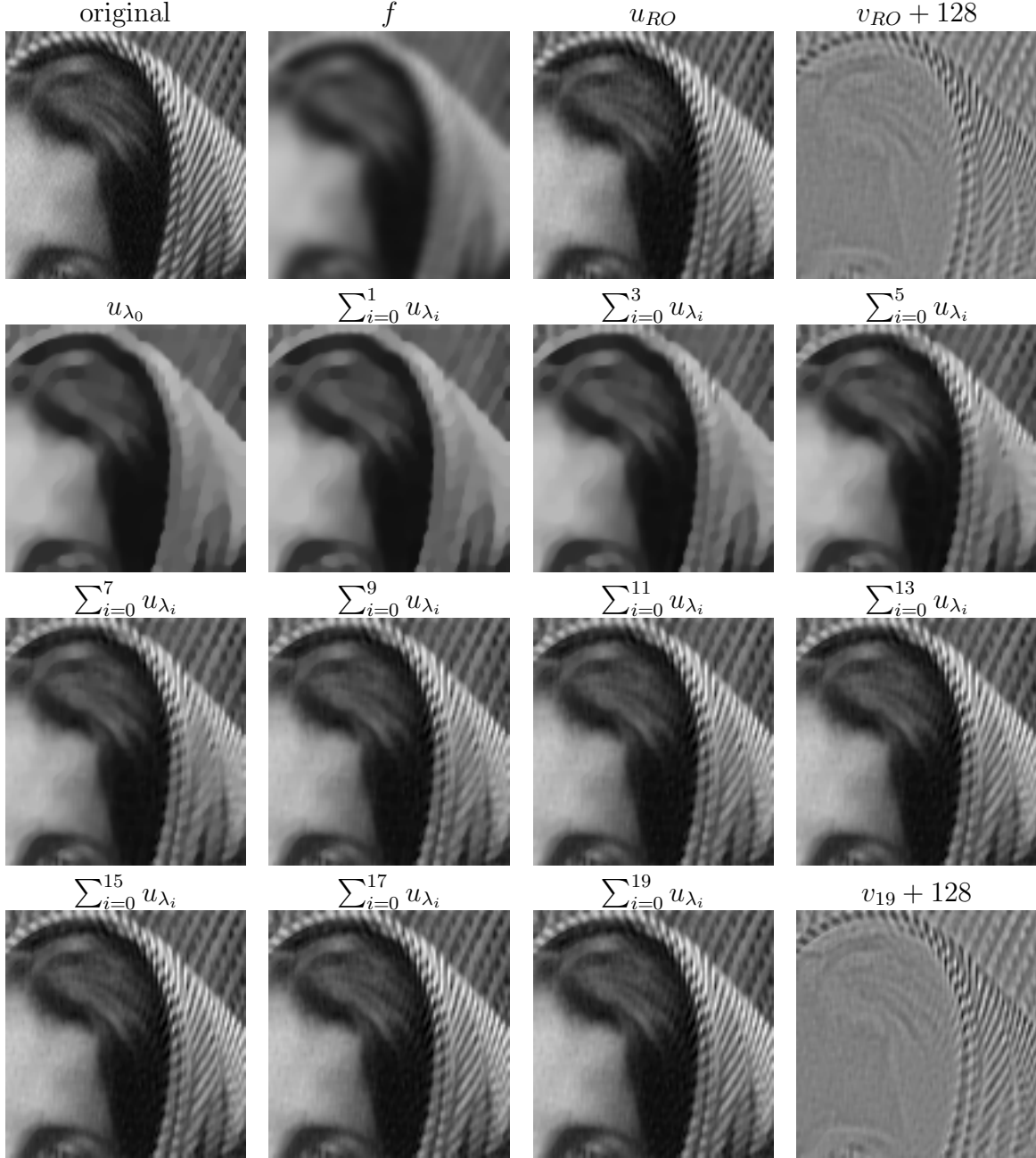
Figure 2.1: 1st row, from left to right: original test image, its blurry version f, and Rudin-Osher restoration u_{RO}, $v_{RO} = f - u_{RO} + 128$, rmse=0.1066. RO parameters: $\lambda = 2000$, $h = 1$, $\triangle t = 0.1$. 2nd to last rows, left to right: hierarchical recovery from the same blurry initial image f using $k = 19$ steps, shown every other one, and residual at step 19. Parameters: $\lambda_0 = 1$, $\triangle t = 0.1$, $h = 1$, and $\lambda_k = 2^k \lambda_0$. Final root mean square error, rmse= 0.0922.

3.1. **Characterization of** (SBV, L^2)**-minimizers.** We proceed along the lines of our analysis of general $\mathcal{Q}$-functionals in section 2.1. We begin with a general characterization of $\mathcal{AT}_\rho$-minimizers as extremal pairs. To this end, we introduce the *weighted spaces* for given $w \in$
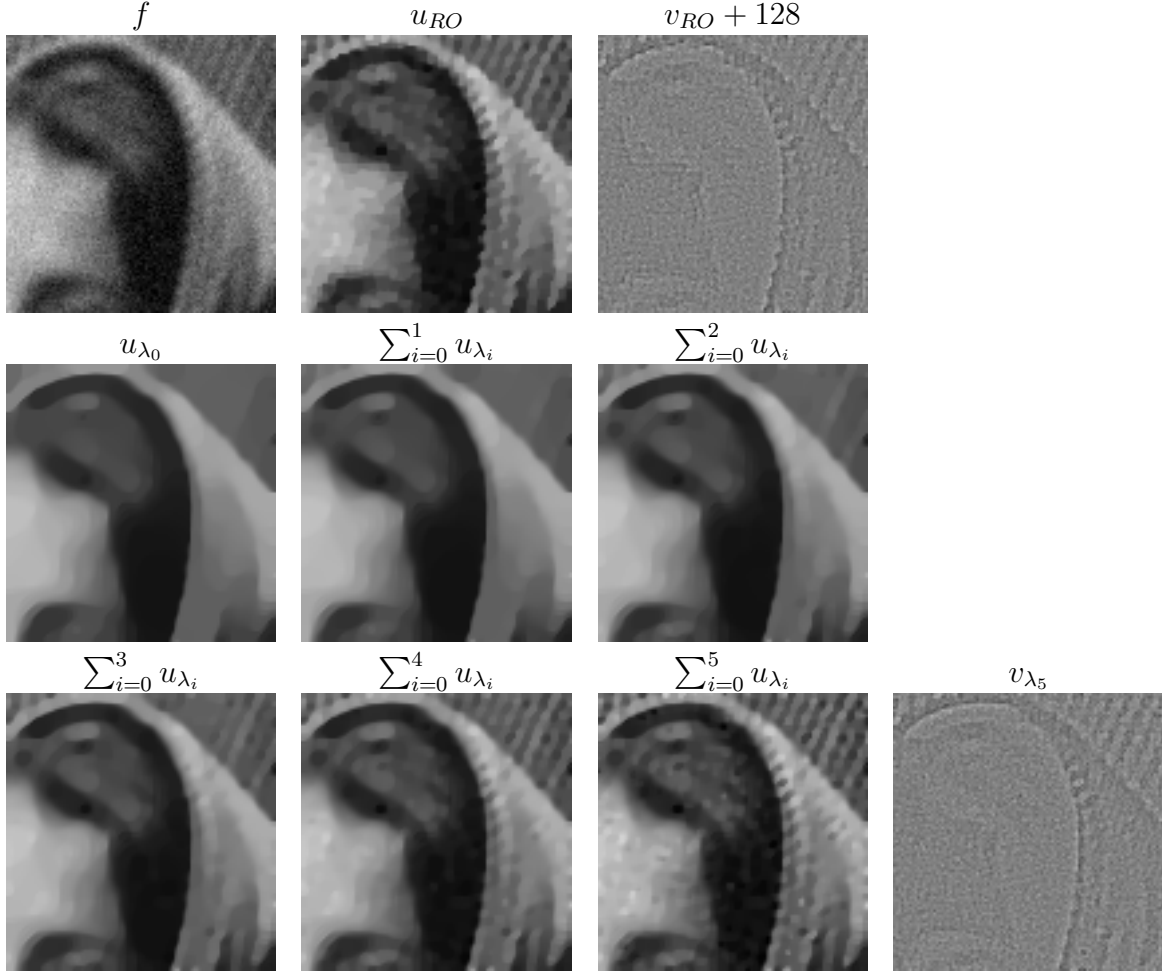
Figure 2.2: 1st row, from left to right: blurry-noisy version f, and Rudin-Osher restoration u_{RO}, $v_{RO} = f - u_{RO} + 128$, rmse =0.2028. RO parameters $\lambda = 0.5$, $h = 1$, $\triangle t = 0.025$. The other rows, left to right: the hierarchical recovery of u from the same blurry-noisy initial image f using 6 steps. Parameters: $\lambda_0 = 0.02$, $\triangle t = 0.025$, $h = 1$, and $\lambda_k = 2^k \lambda_0$, rmse= 0.2011.

$H^1(\Omega)$,

$$|h|_{H_w^1(\Omega)} := \sqrt{\int_\Omega (w^2 + \rho\epsilon_\rho)|\nabla h|^2 dx},$$

and we let

$$\|f\|_{H_w^{-1}(\Omega)} := sup_{h \in H_w^1(\Omega)} \langle f, h \rangle / |h|_{H_w^1(\Omega)}$$

denote the dual norm. We have now the following characterization of the minimizers u, w of the $\mathcal{AT}_\rho$-energy.

Theorem 3.1. *If $u, w \in [0, 1]$ are minimizers of the $\mathcal{AT}_\rho$-energy, then*

(3.1) $$\|f - u\|_{H_w^{-1}(\Omega)} = \frac{\mu}{2\lambda}, \quad \text{and} \quad \langle f - u, u \rangle = \frac{\mu}{2\lambda}|u|_{H_w^1(\Omega)}.$$

Proof. Let $[u, w]$ be a minimizing pair. Considering the variation of $\mathcal{AT}_\rho$ only with respect to u, we find that for any $h \in H^1(\Omega)$, we have

$$\mu|u|_{H^1_w(\Omega)} + \lambda\|f - u\|^2_{L^2(\Omega)} \leq \mu|u + \epsilon h|_{H^1_w(\Omega)} + \lambda\|f - u - \epsilon h\|^2_{L^2(\Omega)}.$$

The triangle inequality yields the "first variation" (or more precisely, sub-differential)

(3.2) $$2\lambda\epsilon\langle f - u, h\rangle \leq \lambda\epsilon^2\|h\|^2_{L^2(\Omega)} + \mu\epsilon|h|_{H^1_w(\Omega)}.$$

For $\epsilon > 0$ this gives $2\lambda\langle f - u, h\rangle \leq \lambda\epsilon\|h\|^2_{L^2(\Omega)} + \mu|h|_{H^1_w(\Omega)}$, and as $\epsilon \to 0_+$, we deduce that for all $h \in H^1_w(\Omega)$,

$$\frac{\langle f - u, h\rangle}{|h|_{H^1_w(\Omega)}} \leq \frac{\mu}{2\lambda},$$

or

(3.3) $$\|f - u\|_{H^{-1}_w(\Omega)} \leq \frac{\mu}{2\lambda}.$$

Now, if we set $h = u$ in (3.2), we find

$$2\lambda\epsilon\langle f - u, u\rangle \leq \lambda\epsilon^2\|u\|^2_{L^2(\Omega)} + \mu\epsilon|u|_{H^1_w(\Omega)}.$$

Again, first dividing by $\epsilon < 0$ and let $\epsilon \uparrow 0_-$, we obtain

(3.4) $$\langle f - u, u\rangle \geq \frac{\mu}{2\lambda}|u|_{H^1_w(\Omega)}.$$

Combining (3.3) and (3.4) we find,

$$\frac{\mu}{2\lambda}|u|_{H^1_w(\Omega)} \leq \langle f - u, u\rangle \leq \|f - u\|_{H^{-1}_w(\Omega)}|u|_{H^1_w(\Omega)} \leq \frac{\mu}{2\lambda}|u|_{H^1_w(\Omega)},$$

confirming that the last inequalities are in fact equalities and thus concluding the proof. $\square$

As before, consult corollary 2.2, the fact that the image f contains too much texture is linked to a trvivial $\mathcal{AT}_\rho$-minimizer. One part of this link is the content of the following theorem.

Theorem 3.2. *If $u \equiv 0$ and $w \equiv 1$ are minimizers of the $\mathcal{AT}_\rho$-energy, then*

(3.5) $$\|f\|_{H^{-1}(\Omega)} \leq \frac{\mu\sqrt{1 + \rho\epsilon_\rho}}{2\lambda}.$$

Proof. If $[u, w] \equiv [0, 1]$ is a $\mathcal{AT}_\rho$-minimizer then for any $h \in H^1(\Omega)$, that

$$\lambda\|f\|^2_{L^2(\Omega)} \leq \epsilon\mu\sqrt{\int_\Omega (1 + \rho\epsilon_\rho)|\nabla h|^2 dx} + \lambda\|f - \epsilon h\|^2_{L^2(\Omega)},$$

or, after expanding terms,

$$2\epsilon\lambda\langle f, h\rangle \leq \epsilon^2\lambda\|h\|^2_{L^2(\Omega)} + \epsilon\mu\sqrt{\int_\Omega (1 + \rho\epsilon_\rho)|\nabla h|^2 dx}.$$

As $\epsilon \downarrow 0_+$, we deduce that

$$2\lambda\langle f, h\rangle \leq \mu\sqrt{\int_\Omega (1 + \rho\epsilon_\rho)|\nabla h|^2 dx},$$

and (3.5) follows. $\square$

Equipped with the characterization of $\mathcal{AT}_\rho$-minimizers, we turn to analyze the corresponding hierarchical decomposition.

Theorem 3.3.
(i)[Hierarchical decomposition]. Consider $f \in L^2(\Omega)$. Then f admits the following hierarchical decomposition

$$(3.6a) \qquad\qquad f \cong \sum_{j=0}^{\infty} u_j,$$

where $\cong$ is interpreted as weak H_w^{-1}-convergence of the residuals

$$(3.6b) \qquad\qquad \|f - \sum_{j=0}^{k} u_j\|_{H_{w_k}^{-1}(\Omega)} = \frac{\mu}{\lambda_0 2^{k+1}}.$$

Here, w_k is computed recursively as the weighting minimizer of $\mathcal{AT}_\rho(f - \sum_{j=0}^{k} u_j, \lambda_0 2^j)$.
(ii)[Energy decomposition]. The following 'energy' estimate holds

$$(3.7) \qquad\qquad \sum_{j=0}^{\infty} \left[\frac{1}{\lambda_j} |u_j|_{H_{w_j}^1(\Omega)} + \|u_j\|_{L^2}^2 \right] \le \|f\|_{L^2}^2, \quad \lambda_j := \lambda_0 2^j.$$

Moreover, if f is sufficiently smooth then equality holds in (3.7).

Proof. The first statement, (3.6), follows from the basic hierarchical expansion, $f = \sum_0^k u_j + v_k$, while noting that $\|v_k\|_{H_{w_k}^{-1}(\Omega)} = \mu/2\lambda_k$. For the second statement, (3.7), we begin by squaring the basic refinement step, $u_{j+1} + v_{j+1} = v_j$,

$$(3.8) \qquad \|v_{j+1}\|_{L^2}^2 + \|u_{j+1}\|_{L^2}^2 + 2\langle u_{j+1}, v_{j+1}\rangle = \|v_j\|_{L^2}^2, \quad j = -1, 0, 1, \ldots.$$

Observe that the last equality holds for $j = -1$ with v_{-1} interpreted as $v_{-1} := f$. We recall that $[u_{j+1}, v_{j+1}]$ is a minimizing pair for $\mathcal{AT}_\rho(v_j, \lambda_{j+1})$ and hence, by (3.1),

$$2\langle u_{j+1}, v_{j+1}\rangle = \frac{\mu}{\lambda_{j+1}} |u_{j+1}|_{H_{w_{j+1}}^1(\Omega)},$$

yielding $\frac{\mu}{\lambda_{j+1}} |u_{j+1}|_{H_{w_{j+1}}^1(\Omega)} + \|u_{j+1}\|_{L^2}^2 = \|v_j\|_{L^2}^2 - \|v_{j+1}\|_{L^2}^2$. We sum up obtaining

$$
\begin{aligned}
\sum_{j=0}^{k} \left[\frac{\mu}{\lambda_j} |u_j|_{H_{w_j}^1(\Omega)} + \|u_j\|_{L^2}^2 \right] &= \sum_{j=-1}^{k-1} \left[\frac{\mu}{\lambda_{j+1}} |u_{j+1}|_{H_{w_{j+1}}^1(\Omega)} + \|u_{j+1}\|_{L^2}^2 \right] = \\
&= \|v_{-1}\|_{L^2}^2 - \|v_k\|_{L^2}^2 = \|f\|_{L^2}^2 - \|v_k\|_{L^2}^2 \le \|f\|_2^2.
\end{aligned}
$$

$\square$

3.2. Discretization of Euler-Lagrange and numerical results.
We consider the (original) Ambrosio-Tortorelli functional [AT92] (neglecting the term $\rho\epsilon_\rho$)

$$\inf_{\{u,v,w \mid u+v=f\}} \left\{ \mu \int_\Omega (w^2) |\nabla u|^2 dx + \rho \|\nabla w\|_{L^2(\Omega)}^2 + \frac{\|w - 1\|_{L^2(\Omega)}^2}{4\rho} + \lambda \|v\|_{L^2(\Omega)}^2 \right\}.$$

The associated Euler-Lagrange equations are

$$
\begin{cases}
\lambda u_\lambda - \mu \nabla(w_\lambda^2 \nabla u_\lambda) = \lambda f, \\
\\
-\triangle w_\lambda + \dfrac{1 + 4\mu\rho|\nabla u_\lambda|^2}{4\rho^2}\left(w_\lambda - \dfrac{1}{1 + 4\mu\rho|\nabla u_\lambda|^2}\right) = 0.
\end{cases}
$$

We construct the hierarchical decomposition in the same manner as before, so $f \sim \sum_{j=0}^{k} u_j$ with the additional feature that the accumulated w_j's give the set of edges of the image u.

Discretization of the Euler-Lagrange equations yields

$$
\begin{cases}
\lambda f_{i,j} = \lambda u_{i,j} - \mu D_-^x(w_{i,j}^2 D_+^x u_{i,j}) - \mu D_-^y(w_{i,j}^2 D_+^y u_{i,j}), \\
\\
D_-^x D_+^x w_{i,j} + D_-^y D_+^y w_{i,j} = \dfrac{1 + 4\mu\rho[(D_0^x u_{i,j})^2 + (D_0^y u_{i,j})^2]}{4\rho^2}\left(w_{i,j} - \dfrac{1}{1 + 4\mu\rho[(D_0^x u_{i,j})^2 + (D_0^y u_{i,j})^2]}\right).
\end{cases}
$$

Using the notation

$$
\begin{aligned}
C_1 : &= \lambda + \frac{\mu}{h^2}(2w_{i,j}^2 + w_{i-1,j}^2 + w_{i,j-1}^2) \\
C_2 : &= 1 + 4\mu\rho\sqrt{\left(\frac{u_{i+1,j} - u_{i-1,j}}{2h}\right)^2 + \left(\frac{u_{i,j+1} - u_{i,j-1}}{2h}\right)^2 + \frac{16\rho^2}{h^2}},
\end{aligned}
$$

we have

$$
\begin{cases}
u_{i,j} = \dfrac{1}{C_1}\left[\lambda f_{i,j} + \dfrac{\mu}{h^2}(w_{i,j}^2(u_{i+1,j} + u_{i,j+1}) + w_{i-1,j}^2 u_{i-1,j} + w_{i,j-1}^2 u_{i,j-1})\right] \\
\\
w_{i,j} = \dfrac{1}{C_2}\left[1 + \dfrac{4\rho^2}{h^2}(w_{i+1,j} + w_{i-1,j} + w_{i,j+1} + w_{i,j-1})\right].
\end{cases}
$$

In order to minimize the grid effect, we alternate the above scheme with the following one, obtained by substituting D_+ for D_- (and vice-versa) in the discretization of the above Euler-Lagrange equation:

$$
\begin{aligned}
D_1 : &= \lambda + \frac{\mu}{h^2}(2w_{i,j}^2 + w_{i+1,j}^2 + w_{i,j+1}^2) \\
D_2 : &= 1 + 4\mu\rho\sqrt{\left(\frac{u_{i+1,j} - u_{i-1,j}}{2h}\right)^2 + \left(\frac{u_{i,j+1} - u_{i,j-1}}{2h}\right)^2 + \frac{16\rho^2}{h^2}},
\end{aligned}
$$

yielding

$$
\begin{cases}
u_{i,j} = \dfrac{1}{D_1}\left[\lambda f_{i,j} + \dfrac{\mu}{h^2}(w_{i,j}^2(u_{i-1,j} + u_{i,j-1}) + w_{i+1,j}^2 u_{i+1,j} + w_{i,j+1}^2 u_{i,j+1})\right] \\
\\
w_{i,j} = \dfrac{1}{D_2}\left[1 + \dfrac{4\rho^2}{h^2}(w_{i+1,j} + w_{i-1,j} + w_{i,j+1} + w_{i,j-1})\right].
\end{cases}
$$

In figures 3.3 and 3.4 we demonstrate the hierarchical $\mathcal{AT}_\rho$-decompositions. We can clearly see that we converge to the desired image as well as obtaining a hierarchical representation, $\sum 2^{-j} w_j$, of the contours of the image f.
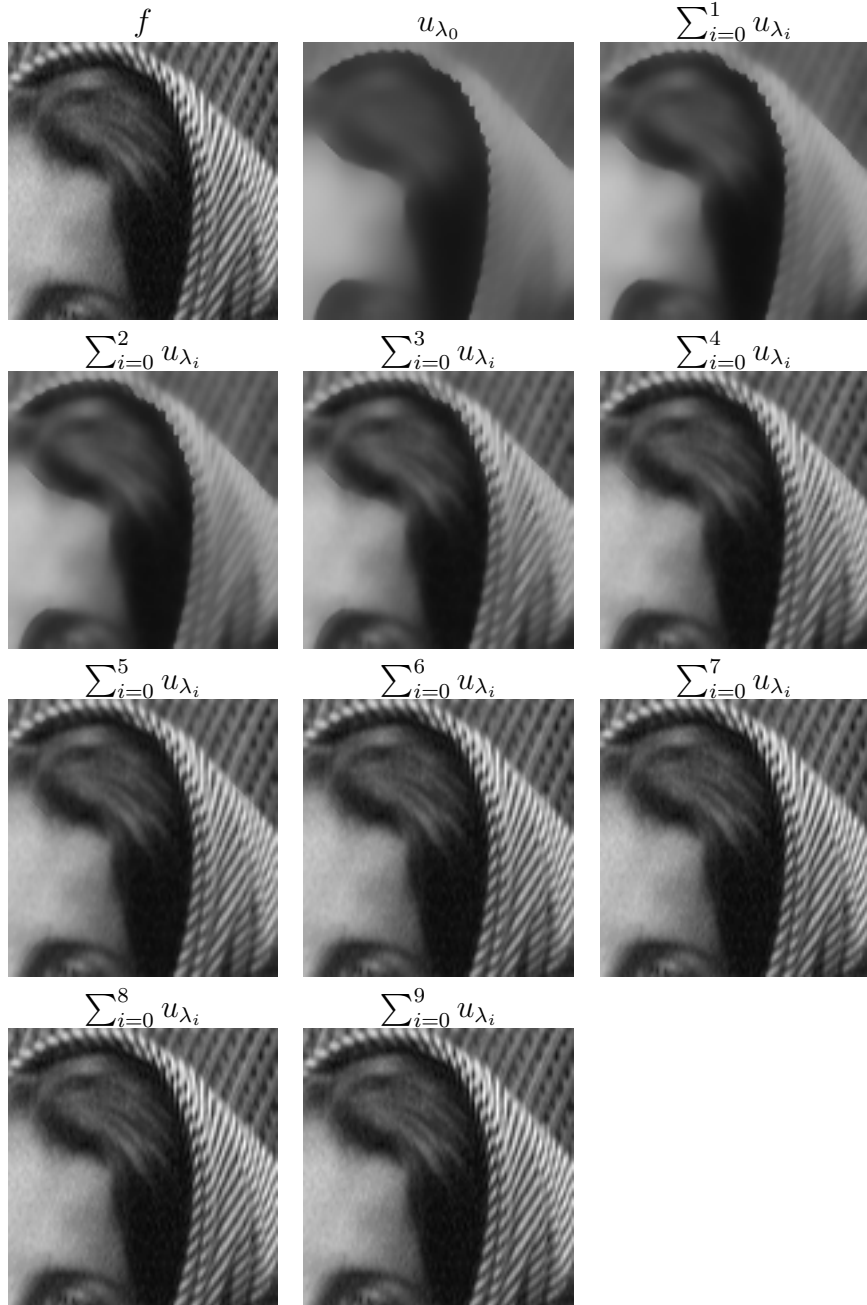
Figure 3.3: The sum of the u_i's using the Ambrosio-Tortorelli approximation of the image of a woman, using 10 steps. Parameters: $\lambda_0 = .25$, $\mu = 5$, $\rho = .0002$, and $\lambda_k = 2^k \lambda_0$.

4. Hierarchical decomposition of images with multiplicative noise

Following [RO94], [RLO03], we consider a *multiplicative* degradation model where we are given an image $f = u \cdot v$, with $u > 0$ being the original image and v models the multiplicative noise, normalized such that $\int_\Omega v(x,y)dxdy = 1$, where for simplicity, we assume that $|\Omega| = 1$. Let u_λ be the minimizer of the corresponding total variation functional in the multiplicative
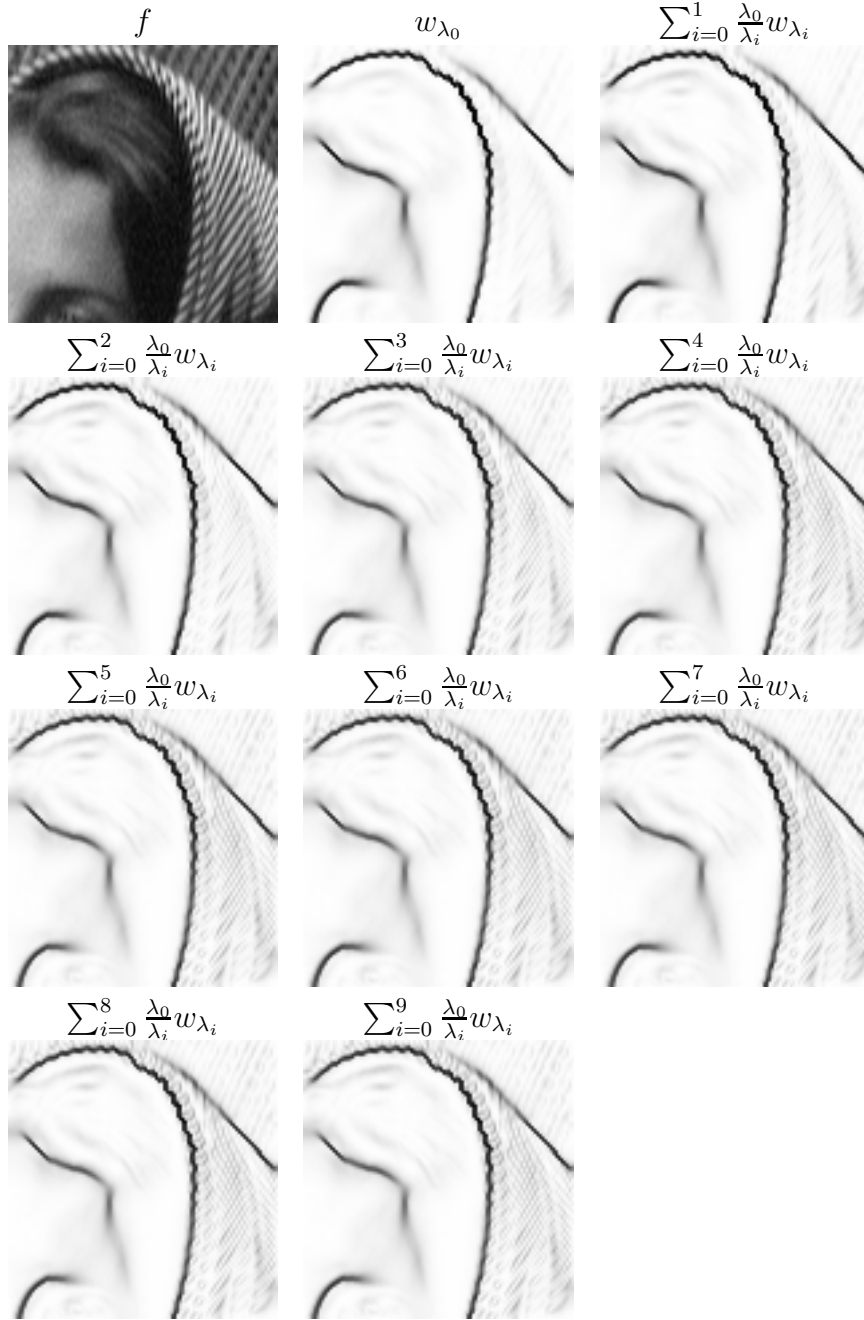
Figure 3.4: The weighted sum of the w_i's using the Ambrosio-Tortorelli approximation of the image of a woman, using 10 steps. Parameters: $\lambda_0 = .25$, $\mu = 5$, $\rho = .0002$, and $\lambda_k = 2^k \lambda_0$.

case [RO94], [CL95] (in a simplified form, without a mean constraint),

$$(4.1) \qquad \mathcal{M}(f, \lambda; BV, L^2) := \inf_{u \in BV_+(\Omega)} \left\{ \lambda \left\| \frac{f}{u} - 1 \right\|_{L^2(\Omega)}^2 + |u|_{BV(\Omega)} \right\}.$$

If $f > 0$ a.e. in Ω, then (4.1) has at least a minimizer $u \geq 0$ [CL95].

Setting $v_\lambda := \frac{f}{u_\lambda}$ we end up with the one scale decomposition $f = u_\lambda v_\lambda$. We construct the hierarchical decomposition as in TNV, except that sums and differences are replaced by
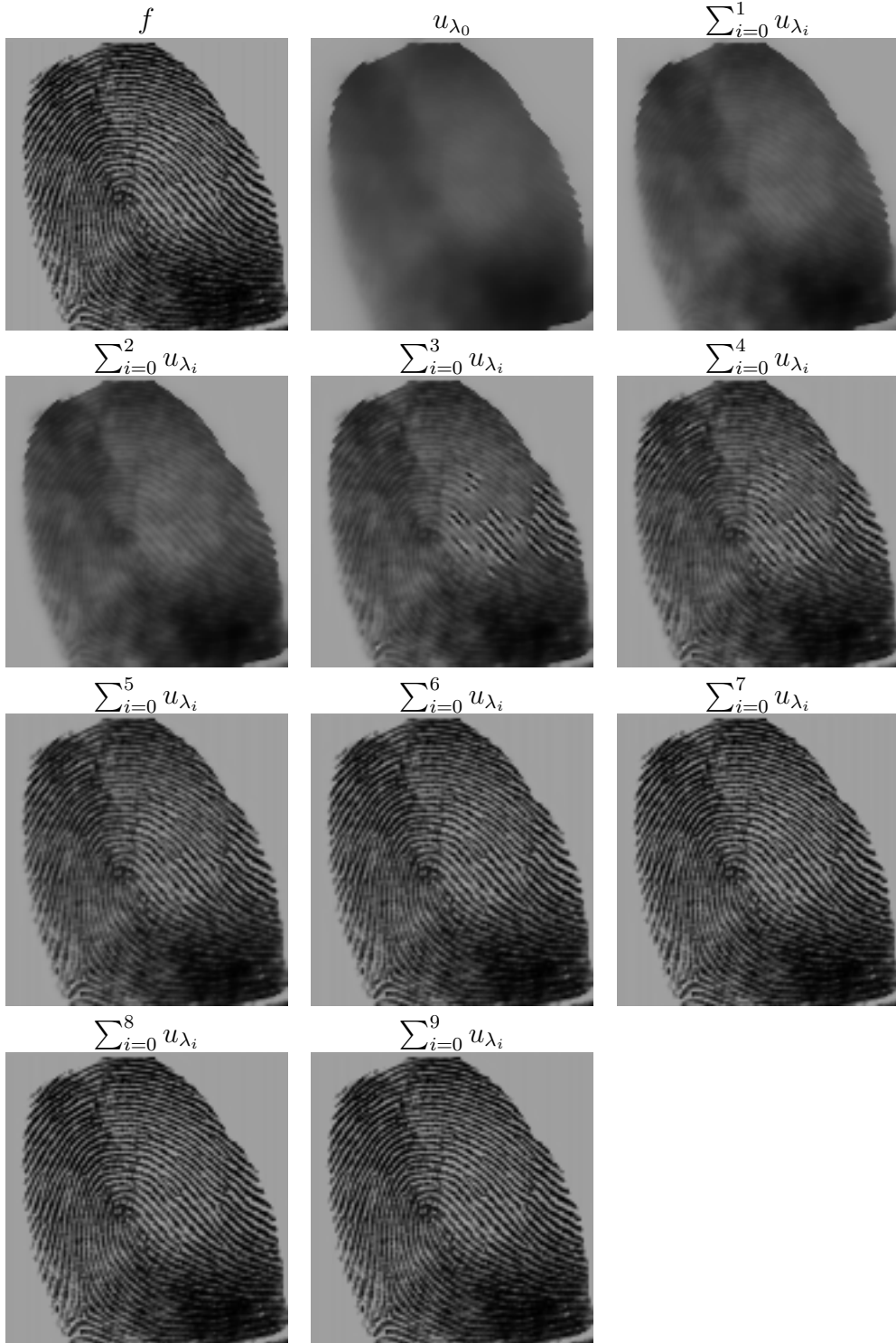
Figure 3.5: The sum of the u_i's using the Ambrosio-Tortorelli approximation on the image of a fingerprint. Parameters: $\mu = 5$, $\rho = .0002$, $\lambda_0 = .25$, $k = 10$, and $\lambda_k = 2^k \lambda_0$.

Figure 3.6: The sum of the w_i's using the Ambrosio-Tortorelli approximation on the image of a fingerprint. Parameters: $\mu = 5$, $\rho = .0002$, $\lambda_0 = .25$, $k = 10$, and $\lambda_k = 2^k \lambda_0$.
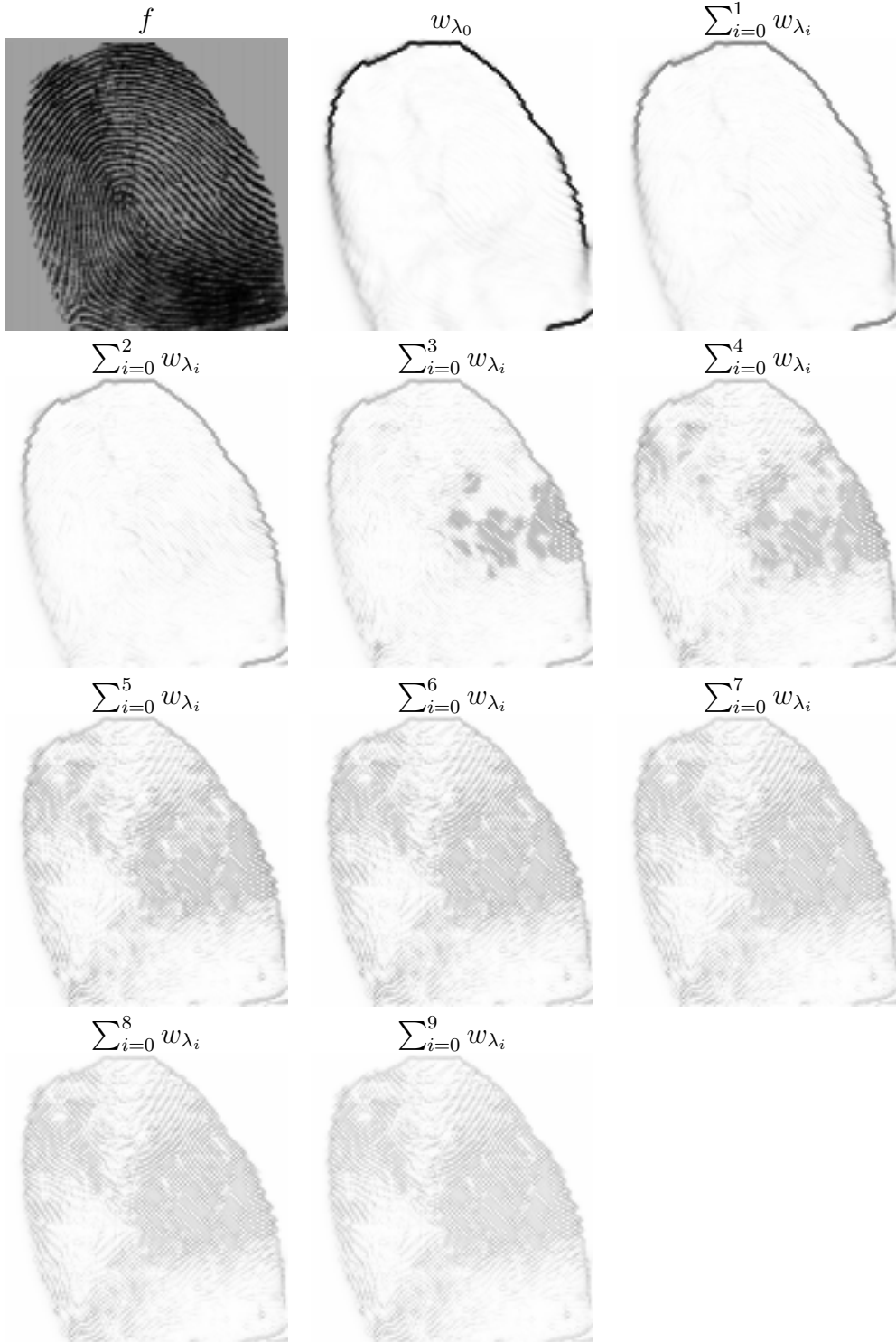
products and quotients. Thus, the iterative step at scale $\lambda_j = \lambda_0 2^j$ reads, $v_j = u_{j+1}v_{j+1}$, leading to the multiplicative hierarchical decomposition

$$f = u_0 u_1 ... u_k \times v_k, \qquad \lambda_j = \lambda_0 2^j.$$

4.1. Characterization of $\mathcal{M}$-minimizers.

We begin with characterization of the $\mathcal{M}(f, \lambda; BV, L^2)$-minimizers. We show that $[u_\lambda, v_\lambda]$ is a minimizer if it is an extremal pair, properly interpreted in terms of the dual BV norm $\|\cdot\|_*$ (consult [Me02, Definition 10]).

Theorem 4.1. *If u is a minimizer of (4.1) then*

$$(4.2) \qquad |u|_{BV(\Omega)} \cdot \left\| \frac{f}{u^2}\left(\frac{f}{u} - 1\right) \right\|_* = \frac{|u|_{BV(\Omega)}}{2\lambda}, \quad and \quad \left\langle \frac{f}{u^2}\left(\frac{f}{u} - 1\right), u \right\rangle = \frac{1}{2\lambda}|u|_{BV(\Omega)}.$$

Proof. Let

$$(4.3) \qquad\qquad g(\epsilon) := \left(\frac{f}{u + \epsilon h} - 1\right)^2.$$

Taylor's expansion gives $g(\epsilon) = \left(\frac{f}{u} - 1\right)^2 - 2\epsilon\left(\frac{f}{u} - 1\right)\frac{fh}{u^2} + \frac{\epsilon^2}{2}g''(\epsilon_\xi)$ and hence

$$\lambda\left\|\frac{f}{u+\epsilon h} - 1\right\|_{L^2(\Omega)}^2 \leq \lambda\left\|\frac{f}{u} - 1\right\|_{L^2(\Omega)}^2 - 2\lambda\epsilon\left\langle \frac{f}{u} - 1, \frac{fh}{u^2}\right\rangle + \lambda\frac{\epsilon^2}{2}\max_x|g''(x)|.$$

This inequality, together with the fact that u is an $\mathcal{M}$-minimizer satisfies for all $h \in BV(\Omega)$,

$$\lambda\left\|\frac{f}{u} - 1\right\|_{L^2(\Omega)}^2 + |u|_{BV(\Omega)} \leq \lambda\left\|\frac{f}{u+\epsilon h} - 1\right\|_{L^2(\Omega)}^2 + |u + \epsilon h|_{BV(\Omega)},$$

imply that

$$(4.4) \qquad\qquad 2\lambda\epsilon\left\langle \frac{f}{u} - 1, \frac{fh}{u^2}\right\rangle \leq \epsilon|h|_{BV(\Omega)} + \lambda\frac{\epsilon^2}{2}\max_x|g''(x)|.$$

Dividing by ϵ and letting $\epsilon \downarrow 0_+$ (while noticing that $\lim_{\epsilon \to 0}\frac{\epsilon^2}{2}\max_x|g''(x)| = 0$), yield that for any $h \in BV(\Omega)$,

$$2\lambda\left\langle \frac{f}{u^2}\left(\frac{f}{u} - 1\right), h\right\rangle \leq |h|_{BV(\Omega)};$$

thus

$$(4.5) \qquad\qquad \left\|\frac{f}{u^2}\left(\frac{f}{u} - 1\right)\right\|_* \leq \frac{1}{2\lambda}.$$

Now let $h = u$ in (4.4). Then, dividing by $\epsilon < 0$, and letting $\epsilon \uparrow 0_-$, we obtain

$$(4.6) \qquad\qquad \left\langle \frac{f}{u^2}\left(\frac{f}{u} - 1\right), u\right\rangle \geq \frac{|u|_{BV(\Omega)}}{2\lambda}.$$

Combining (4.5) and (4.6) concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad \square$

As before, theorem 4.1 could be refined, depending on the amount of texture present in f. For example, compared with corollary 2.2, we have that f consists mostly of texture, if and only if its $\mathcal{M}$-minimizer is the trivial one which, in the multiplicative case, is given by $u \equiv 1$. The following theorem confirms one-side of this implication.

Theorem 4.2. *If $u \equiv 1$ is a minimizer of (4.1), then $\|f(f - 1)\|_* \leq \frac{1}{2\lambda}$.*

Proof. Since $u \equiv 1$ is assumed to be $\mathcal{M}$-minimizer, we have for any $h \in BV(\Omega)$

$$\lambda \left\| \frac{f}{1 + \epsilon h} - 1 \right\|_{L^2(\Omega)}^2 + |1 + \epsilon h|_{BV(\Omega)} \geq \lambda \| f - 1 \|_{L^2(\Omega)}^2.$$

Revisiting $g(\epsilon)$ in (4.3) with $u \equiv 1$, we have

$$g(\epsilon) := \left(\frac{f}{1 + \epsilon h} - 1 \right)^2 = (f - 1)^2 - 2\epsilon(f - 1)fh + \frac{\epsilon^2}{2} g''(\epsilon_\xi).$$

The last two relations yield

$$(4.7) \qquad \lambda \| f - 1 \|_{L^2(\Omega)}^2 \quad - \quad 2\lambda\epsilon \langle h, f(f - 1) \rangle + \lambda\epsilon^2 \int_\Omega \frac{h^2}{2} g''(1) dx$$

$$(4.8) \qquad + \lambda \int_\Omega \mathcal{O}((\epsilon h)^3) dx + ... + \epsilon |h|_{BV(\Omega)} \geq \lambda \| f - 1 \|_{L^2(\Omega)}^2.$$

Divide by ϵ and let $\epsilon \downarrow 0_+$, to obtain that for all $h \in BV(\Omega)$

$$|h|_{BV(\Omega)} \geq 2\lambda \langle h, f(f - 1) \rangle,$$

which means that $\| f(f - 1) \|_* \leq \frac{1}{2\lambda}$. $\qquad \square$

4.2. Discretization of Euler-Lagrange and numerical results.

Formally minimizing $\mathcal{M}(f, \lambda)$ with respect to u yields the following associated Euler-Lagrange equation:

$$\left(\frac{f}{u_\lambda} - 1 \right) \cdot \left(-\frac{f}{u_\lambda^2} \right) = \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{|\nabla u_\lambda|} \right).$$

When working on a bounded domain Ω, we augment the Euler-Lagrange equations by the following Neumann boundary condition:

$$\frac{\partial u_\lambda}{\partial n}_{|\partial \Omega} = 0.$$

The hierarchical decomposition, $f \sim \prod_{j=0}^k u_j$, is obtained, in which the u_j's are constructed as (approximate) solutions of the recursive relation governed by the Euler-Lagrange equation.

To discretize the Euler-Lagrange equation, we begin by regularization of $\mathcal{M}(f, \lambda)$ to avoid the singularity when $|\nabla u_\lambda| = 0$. So, we have

$$\mathcal{M}^\epsilon(f, \lambda) := \inf_{u \in BV} \left\{ \lambda \left\| \frac{f}{u} - 1 \right\|_{L^2(\Omega)}^2 + \int_\Omega \sqrt{\epsilon^2 + |\nabla u|^2} dx dy \right\}.$$

This yields the associated Euler-Lagrange equations:

$$(4.9) \qquad -\frac{f^2}{u_\lambda^3} + \frac{f}{u_\lambda^2} = \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u_\lambda}{\sqrt{\epsilon^2 + |\nabla u_\lambda|^2}} \right),$$

that we solve by a dynamic scheme $(x, y, t) \mapsto u(x, y, t)$:

$$(4.10) \qquad \frac{\partial u}{\partial t} = \frac{f^2}{u^3} - \frac{f}{u^2} + \frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u}{\sqrt{\epsilon^2 + |\nabla u|^2}} \right), \quad u(x, y, 0) = f(x, y).$$
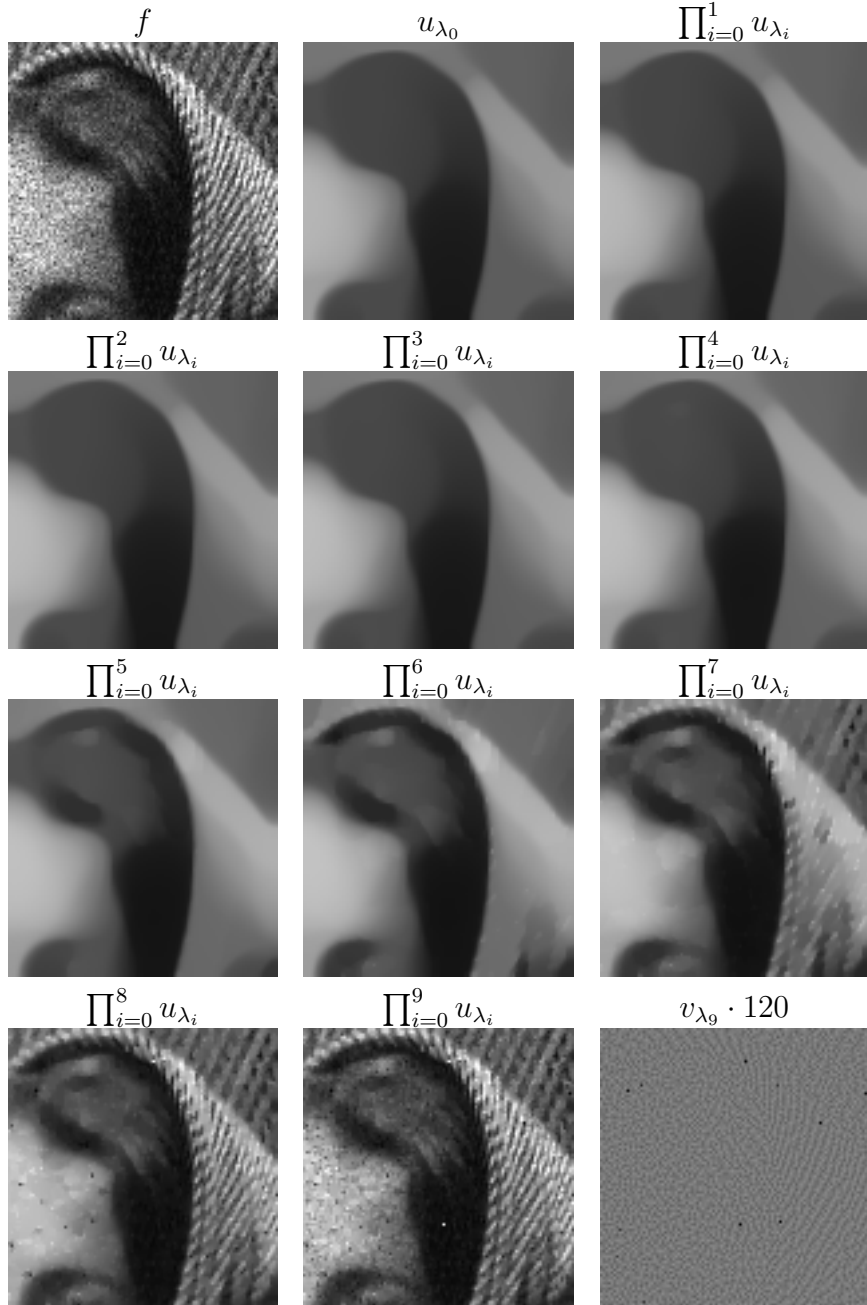
f u_{λ_0} $\prod_{i=0}^{1} u_{\lambda_i}$

$\prod_{i=0}^{2} u_{\lambda_i}$ $\prod_{i=0}^{3} u_{\lambda_i}$ $\prod_{i=0}^{4} u_{\lambda_i}$

$\prod_{i=0}^{5} u_{\lambda_i}$ $\prod_{i=0}^{6} u_{\lambda_i}$ $\prod_{i=0}^{7} u_{\lambda_i}$

$\prod_{i=0}^{8} u_{\lambda_i}$ $\prod_{i=0}^{9} u_{\lambda_i}$ $v_{\lambda_9} \cdot 120$

Figure 4.7: The recovery of u given an initial image of a woman with multiplicative noise, for 10 steps. Parameters: $\lambda_0 = .02$, and $\lambda_k = 2^k \lambda_0$.

Let $u_{i,j}^n \approx u(x_i, y_j, n \triangle t)$. The discretization that we have used is a linearized semi-implicit scheme:

$$
\begin{aligned}
\frac{u_{i,j}^{n+1} - u_{i,j}^n}{\triangle t} &= \frac{f_{i,j}^2}{(u_{i,j}^n)^3} - \frac{f_{i,j}}{(u_{i,j}^n)^2} \\
&+ \frac{1}{2\lambda h^2}(c_E u_{i+1,j}^n + c_W u_{i-1,j}^n + c_S u_{i,j+1}^n + c_N u_{i,j-1}^n) \\
&- \frac{1}{2\lambda h^2}(c_E + c_W + c_S + c_N)u_{i,j}^{n+1},
\end{aligned}
$$

or, solving for $u_{i,j}^{n+1}$,

$$u_{i,j}^{n+1} \;=\; \left(\frac{1}{1 + \frac{\triangle t}{2\lambda h^2}(c_E + c_W + c_S + c_N)}\right) \cdot \left[u_{i,j}^n + \triangle t \frac{f_{i,j}^2}{(u_{i,j}^n)^3} - \triangle t \frac{f_{i,j}}{(u_{i,j}^n)^2}\right.$$

$$\left. + \frac{\triangle t}{2\lambda h^2}(c_E u_{i+1,j}^n + c_W u_{i-1,j}^n + c_S u_{i,j+1}^n + c_N u_{i,j-1}^n)\right],$$

where the constants c_E, c_W, c_S and c_N are the same as before, (2.19).

We note in passing the issue of *stability*: in order to enable the necessary division by $u \neq 0$, we shift f away from zero, adding a positive constant which is subtracted from the final result.

We demonstrate our hierarchical decomposition to the image f in figure 4.7. We can see that just as in the case with additive noise, we must pay a price for the recovered texture, namely the return of some noise. As in the case with additive noise, using a finer decomposition might give improved results.

REFERENCES

[AV94] R. ACAR AND C. R. VOGEL, *Analysis of Bounded Variation Penalty Methods of Ill-Posed Problems*, Inverse Problems, 10 (1994), pp. 1217-1229.

[AT92] L. AMBROSIO AND V. TORTORELLI, *On the Approximation of Free Discontinuity Problems*, Bollettino U.M.I. (7) 6-B (1992), 105-123.

[BS88] C. BENNET AND R. SHARPLEY, Interpolation of Operators, Academic Press, 1988.

[BL76] J. BERGH AND J. LOFSTROM, Interpolation Spaces, Springer Verlag, 1976.

[ACM04] F. ANDREU-VAILLO, V. CASELLES, AND J.M. MAZÓN, Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, Birkhäuser Verlag, 2004.

[AV97] G. AUBERT AND L. VESE, *A Variational Method in Image Recovery*, SIAM J. Numer. Anal., 34 (1997), pp. 1948-1979.

[ACMS98] A. Z. AVERBUCH, R. R. COIFMAN, F. G. MEYER, AND J. O. STRÖMBERG, *Multi-Layered Image Representation: Application to Image Compression*, in Proceedings of the International Conference on Image Processing (ICIP 1998), Vol. 2, Chicago, IL, 1998, IEEE Computer Society Press, Los Alamitos, CA 1998, pp. 292-296.

[BCM05] A. BAUDES, B. COLL AND J.-M. MOREL, *A review of image denoising algorithms, with a new one*, Multiscale Model. Simul. 4 (2005), no. 2, 490–530.

[CL95] A. CHAMBOLLE AND P.L. LIONS, *Image restoration by constrained total variation minimization and variants*, SPIE Electronic Imaging Proceedings, Vol. 2567, 1995.

[CL97] A. CHAMBOLLE AND P. L. LIONS, *Image Recovery via Total Variation Minimization and Related Problems*, Numer. Math., 76 (1997), pp. 167-188.

[DT84] F. DEMENGEL AND R. TEMAM, *Convex functions of a measure and applications*, Indiana Univ. Mathematics Journal 33(5): 673-709, 1984.

[DL93] R.A. DEVORE AND G.G. LORENTZ, Constructive Approximation, Springer-Verlag, Grundlehren der mathematischen Wissenschaften Vol. 303, 1993.

[EO04] S. ESEDOGLU AND S. OSHER, *Decomposition of images by the anisotropic Rudin-Osher-Fatemi model*, Comm. Pure Appl. Math. LVII 1609-1626 (2004).

[GG84] S. GEMAN AND D. GEMAN, *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Mach. Intell, 6, 721-741, 1984.

[GS00] C. W. Groetsch and O. Scherzer, *Nonstationary iterated Tikhonov-Morozov method and third order differential equations for the evaluation of unbounded operators*, Math. Methods Appl. Sci. 23 p. 1287, 2000.

[GS00] C. W. Groetsch and O. Scherzer, *Inverse scale space theory for inverse probelms*, Lecture Notes In Computer Science; Vol. 2106 Proc. Third International Conference on Scale-Space and Morphology in Computer Vision, 317–325, 2001.

[Jo07] J. Jost, Partial Differential Equations, Springer, 2007.

[Kr07] N. Kruglyak, *The K-functional and Calderón-Zygmund type decompositions*, in "Interpolation Theory and Applications" (L. De Carli and M . Milman, eds), Contemporary Mathematics, Amer. Math. Soc., 2007.

[LPSX06] D. Levy, D. Paquin, E. Schreibmann, and L. Xing, *Multiscale Image Registration, Mathematical Biosciences and Engineering*, 3, 2006, pp. 389-418.

[MAC02] F. G. Meyer, A. Z. Averbuch, and R. R. Coifman, *Multilayered Image Representation: Application to Image Compression*, IEEE Trans. Image Process., 11 (2002), pp. 1072-1080.

[Me02] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, University Lecture Series Volume 22, AMS 2002.

[Mo84] V.A. Morozov, Methods for Solving Incorrectly Posed Problems, Springer Verlag, Heidelberg, 1984.

[Mo93] V.A. Morozov, Regularization Methods for Ill-Posed Problems, Florida, CRC Press, 1993.

[MS89] D. Mumford and J. Shah, *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems*, Communications on Pure and Applied Mathematics, Vol. XLVII(1989), pp. 577-685.

[Ne03] S. Nezzar, *A Multiscale Image Representation Using Hierarchical (BV, L^2) Decompositions*, Ph.D. Thesis, UCLA, May 2003.

[RC] L. Rudin and V. Caselles, *Image Recovery via Multiscale Total Variation*, in Proceedings of the Second European Conference on Image Processing, Palma, Spain, 1995.

[PL07] D. Paquin, D. Levy and L. Xing, *Hybrid Multiscale Landmark and Deformable Registration*, Mathematical Biosciences and Engineering, 4 (2007), pp. 711-737.

[RLO03] L. Rudin, P.-L. Lions, and S. Osher, *Multiplicative Denoising and Deblurring: Theory and Algorithms*, in S. Osher and N. Paragios, Editors, "Geometric Level Set Methods in Imaging, Vision and Graphics", Springer 2003, pp. 103-119.

[Rudin] L. Rudin, *Images, Numerical Analysis of Singularities, and Shock Filters*, Ph.D. Thesis, CS Dept., CalTech, Pasadena, #5250:TR:87; (1987).

[ROF92] L. Rudin, S. Osher and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, 60, pp. 259-268, 1992.

[RO94] L. Rudin and S. Osher, *Total Variation Based Image Restoration with Free Local Constraints*, Proc. IEEE ICIP, Vol. I, pp. 31-35, Austin (Texas) USA, 1994.

[SS04] A. J. Smola and B. Scholkop, *A tutorial on support vector regression*, Stat. and Computing 14, 199-222, 2004.

[TNV04] E. Tadmor, S. Nezzar and L. Vese, *A Multiscale Image Representation Using Hierarchical (BV,L2) Decomposition*, Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, Volume 2, Number 4, pp. 554-579, 2004.

[TA77] A. N. Tikhonov and V. A. Arsenin, Solution of Ill-posed Problems, Winston & Sons, Washington, 1977.

[Va98] V. Vapnik, Statistical Learning Theory, John Wiley and Sons, 1998.

[Ve01] L. Vese, *A Study in the BV Space of a Denoising-Deblurring Variational Problem*, Appl. Math. Optim., 44 (2001), pp. 131-161.

Department of Mathematics, Center of Scientific Computation And Mathematical Modeling (CSCAMM) and Institute for Physical Science and Technology (IPST), University of Maryland, MD 20742.

E-mail address: tadmor@cscamm.umd.edu

Richard Stockton College of New-Jersey, Pomona, NJ 08240

E-mail address: nezzars@stockton.edu

Department of Mathematics, UCLA, Los Angeles, CA 90095.

E-mail address: lvese@math.ucla.edu